

ANALISIS CLUSTER MENGGUNAKAN ALGORITMA K-MEANS CLUSTER UNTUK CULSTERING JENIS PENYAKIT MENULAR PADA PUSKESMAS DI KECAMATAN KOTA TANGERANG

CLUSTER ANALYSIS USING K-MEANS CLUSTER ALGORITHM FOR CLUSTERING TYPES OF TRANSMITTED DISEASE IN HEALTH CENTER IN KECAMATAN KOTA TANGERANG

Aninda Rachman¹⁾, Aden, S.Si., M.Pd.²⁾, dan Yulianti Rusdiana, S.Si., M.Sc³⁾

¹⁾²⁾³⁾ Program Studi Matematika, FMIPA Universitas Pamulang, Kampus 2
Jl. Raya Puspitek No.46, Buaran, Serpong
Email : rachmananinda@gmail.com

ABSTRACT

The research in this paper aims to cluster the types of diseases in the sub-district Puskesmas in the city of Tangerang, to facilitate the handling and improvement of the environment, and to find out which disease is the most dominant in each cluster, during clustering at the sub-district Puskesmas in Tangerang City. The subject of this research is the result of clustering of infectious diseases with the method used is Cluster Analysis using the K-means Cluster Algorithm using calculations using SPSS 22. And the results of this study can find out the type of disease with each Puskesmas according to the sub-district, for which clusters are determined there are four clusters where each cluster is a type of infectious disease itself, namely pulmonary Tuberculosis, pneumonia, dengue, and diarrhea, for each cluster each divided according to its sub-district or Puskesmas. With these results, it can be concluded that the clustering results know which sub-districts have contracted infectious diseases according to the cluster formed.

Keywords : Cluster Analysis, K-means Cluster Algorithm, Types of Infectious Diseases, Puskesmas Tangerang City, SPSS 22.

ABSTRAK

Penelitian ini bertujuan untuk Clustering jenis penyakit pada Puskesmas perkecamatan di Kota Tangerang, untuk memudahkan penanganan dan perbaikan lingkungan, Serta mencari jenis penyakit mana yang paling dominan pada masing-masing klaster, saat pengklasteran pada puskesmas perkecamatan di Kota Tangerang. Subjek penelitian ini adalah hasil dari pengklasteran jenis penyakit menular dengan metode yang digunakan adalah Analisis Klaster menggunakan Algoritma K-means Klaster dengan perhitungan menggunakan Aplikasi SPSS 22. Dan hasil penelitian ini dapat mengetahui jenis penyakit dengan masing-masing puskesmas sesuai kecamatan, untuk itu klaster ditentukan ada empat klaster dimana setiap klaster adalah jenis penyakit menular itu sendiri yaitu TB Paru, Pneumonia, DBD, dan Diare, untuk setiap klaster masing-masing terbagi sesuai kecamatannya atau puskesmasnya. Dengan hasil tersebut dapat disimpulkan bahwa hasil pengklasteran mengetahui kecamatan mana saja yang terjangkau penyakit menular sesuai klaster yang terbentuk.

Kata kunci : Analisis Klaster, Algoritma K-means Klaster, Jenis Penyakit Menular, Puskesmas Kota Tangerang, SPSS 22.

1. PENDAHULUAN

Kota Tangerang memiliki luas 164,55% dengan jumlah kecamatan 13 dan 104 Kelurahan, jumlah penduduk sebanyak 2.093.706 jiwa, dan Kota Tangerang memiliki tingkat kepadatan penduduk sebesar 12.724 jiwa/km². dengan jumlah dan kepadatan penduduk yang cukup padat lingkungan Kota Tangerang pun memiliki penyebaran penyakit dari data 30 Puskesmas yang ada di Kota Tangerang. Jenis Penyebaran penyakit menular diantaranya TB Paru yang memang akan sangat mudah menyebar lewat percikan ludah lewat udara, jika kepadatan penduduk cukup padat sehingga ruang udarapun akan sedikit, sehingga penyebaran lewat udara akan sangat mudah penyebarannya, Pneumonia sama halnya dengan TB Paru penyebaran terjadi lewat percikan air liur lewat udara, DBD penyakit ini tergantung dengan kebersihan lingkungan sekitar dan

penyebarannya lewat gigitan nyamuk aedes aegypti berkembangbiak pada lingkungan kotor untuk itu jika jumlah penduduk banyak semakin banyak juga pemanfaatan lingkungan yang erpakai sehingga lingkungan padat penduduk pun tidak bisa dihindari dari lingkungan kotor, dan Diare biasanya penyebarannya memulai hidup tidak sehat lingkungan kotor sehingga banyak bakteri dan virus yang mengarah diare.

Untuk itu patut diketahui masing-masing daerah memiliki penyebaran jenis penyakit mana yang paling dominan pada puskesmas perkecamatan kota Tangerang, pengklasteran pada jenis penyakit menular dengan metode Algoritma K-means Klaster dapat diketahui jenis penyakit dan daerah penyebarannya. Dengan demikian, masalah yang dapat diidentifikasi adalah Bagaimana Pembentukan Klaster atau Clustering jenis penyakit pada Puskesmas perkecamatan di Kota Tangerang untuk memudahkan penanganan dan perbaikan lingkungan. Serta mencari jenis penyakit mana yang paling dominan pada masing-masing klaster, saat mengklusterkan pada puskesmas perkecamatan di Kota Tangerang. pembatasan masalah dalam penelitian ini yaitu: objek penelitian ini adalah Jenis Penyakit pada Puskesmas setiap Kecamatan di Kota Tangerang; Pengklasteran yang dipengaruhi Jumlah penduduk, kepadatan penduduk, dan jenis penyakit yang di derita pada puskesmas perkecamatan di Kota Tangerang; Subjek penelitian ini adalah hasil dari Pengklasteran Jenis Penyakit pada Puskesmas perkecamatan di Kota Tangerang; Metode yang digunakan adalah Analisis Klaster Menggunakan Algoritma K-means Klaster dengan perhitungan menggunakan Aplikasi SPSS 22.

2. DATA DAN METODE

2.1 Data dan Lokasi

Data yang digunakan dalam penelitian ini merupakan data Sekunder yaitu berupa data yang dikumpulkan dari sumber BPS Kota Tangerang dengan nama Publikasi Katalog Kecamatan dalam Angka melalui web <https://tangerangkota.bps.go.id>. sumber data sudah berbentuk dalam dokumen yang peneliti jadikan sebagai populasi dan sampel. Berikut adalah tabel data yang akan diklasterkan.

Tabel 2.1 data

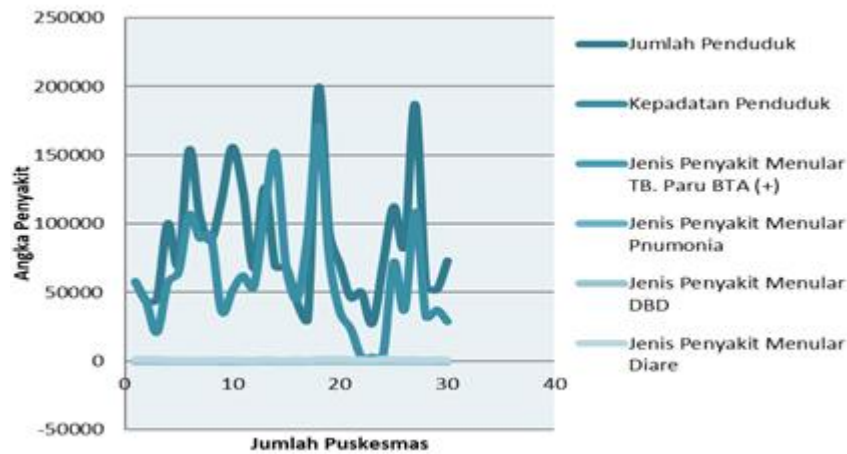
No.	Kecamatan	Puskesmas	Jumlah Penduduk	Kepadatan Penduduk	Jenis Penyakit Menular			
					TB. Paru	Pnumonia	DBD	Diare
1	Batu Ceper	Batu Ceper	57924	57351	31	116	8	1296
2		Poris Gaga	43462	43032	64	180	7	1384
3	Benda	Jurumudi Baru	45049	20555	19	155	1	1504
4		Benda	100059	57118	22	143	3	789
5	Cibodas	Cibodasari	70461	64513	52	260	23	715
6		Baja	153768	106488	97	240	16	420
7	Ciledug	Tajur	104708	89211	46	107	11	643
8		Ciledug	87683	91563	75	142	5	692
9	Cipondoh	Gondrong	117197	36264	11	153	26	983
10		Ketapang	155801	50819	28	294	41	946
11	Jatiuwung	Jatiuwung	124066	62121	39	461	36	543
12	Karang Tengah	Pondok Bahar	67521	53814	28	69	8	408
13		Karang Tengah	126315	108794	51	708	12	1221
14	Karawaci	Pabuaran Tumpeng	69137	151604	40	250	18	734
15		Bugel	68482	69010	15	126	20	308
16		Karawaci Baru	43172	44859	42	221	32	1089
17		Pasar Baru	30407	90704	9	55	9	290
18	Larangan	Larangan Utara	198950	171267	35	163	3	1441
19		Cipadu	94177	76435	70	168	21	1625
20	Neglasari	Neglasari	70424	35658	58	321	5	1666
21		Kedaung Wetan	46128	23457	18	192	8	999
22	Periuk	Periuk Jaya	50136	2766	45	182	0	1031
23		Sangiang Jaya	27103	2657	34	67	9	651
24		Gembor	69581	3434	48	116	16	1762
25	Pinang	Kunciran	112486	71928	23	197	2	1335
26		Kunciran Baru	84262	36984	15	141	10	937
27		Panunggangan	186932	108752	42	180	8	1162
28	Tangerang	Sukasari	54673	32288	36	121	27	474
29		Tanah Tinggi	51704	37207	29	253	22	1329
30		Cikokol	72958	28473	22	162	28	504

2.2 Metode Analisis

Analisis data pengklasteran menggunakan *Non-Hierarchical Cluster* dengan metode penentuan kelompok Algoritma *K-means Cluster*. metode ini justru dimulai dengan terlebih dahulu jumlah klaster yang diinginkan (dua klaster, tiga klaster atau yang lain) sampai terbentuknya klaster.

Dari data di atas akan di cari empat klaster dengan menggunakan aplikasi SPSS 22 dan untuk menguji klaster yang telah ditentukan sudah akurat akan membuat chart yang menunjukkan distribusi normal, matriks similarity, dan korelasi data.

Dari gambar chart di berikut, data yang dipakai yaitu jumlah puskesmas dengan angka penyakit memiliki data berdistribusi normal demikian untuk variabelnya pun memiliki karakteristik-karakteristik yang berbeda tetapi tetap berdistribusi normal.



Gambar 2.1. Gambar Chart Berdistribusi Normal

Ukuran kesamaan atau (similarity) pada data menggunakan kesamaan jarak antara titik pada data puskesmas memiliki kesamaan yang berbeda jadi untuk pengujian metode analisis kluster dengan empat kluster yang akan dibentuk dapat dilakukan.

Tabel 2.2 Korelasi

			Correlations					
Control Variables			Jumlah Penduduk	Kepadatan Penduduk	TB.Paru	Pneumonia	DBD	Diare
-none. ^a	Jumlah Penduduk	Correlation	1.000	.614	.210	.321	.099	.088
		Significance (2-tailed)	.	.000	.266	.083	.601	.642
		df	0	28	28	28	28	28
	Kepadatan Penduduk	Correlation	.614	1.000	.251	.228	-.089	-.116
		Significance (2-tailed)	.000	.	.182	.226	.641	.540
		df	28	0	28	28	28	28
	TB.Paru	Correlation	.210	.251	1.000	.260	-.034	.156
		Significance (2-tailed)	.266	.182	.	.165	.860	.411
		df	28	28	0	28	28	28
	Pneumonia	Correlation	.321	.228	.260	1.000	.267	.176
		Significance (2-tailed)	.083	.226	.165	.	.154	.353
		df	28	28	28	0	28	28
	DBD	Correlation	.099	-.089	-.034	.267	1.000	-.266
		Significance (2-tailed)	.601	.641	.860	.154	.	.155
		df	28	28	28	28	0	28
	Diare	Correlation	.088	-.116	.156	.176	-.266	1.000
		Significance (2-tailed)	.642	.540	.411	.353	.155	.
		df	28	28	28	28	28	0
TB.Paru & Pneumonia & DBD & Diare	Jumlah Penduduk	Correlation	1.000	.614				
		Significance (2-tailed)	.	.001				
		df	0	24				
	Kepadatan Penduduk	Correlation	.614	1.000				
		Significance (2-tailed)	.001	.				
		df	24	0				

a. Cells contain zero-order (Pearson) correlations.

Dari tabel korelasi di atas antara variabel TB Paru, Pneumonia, DBD, Diare dengan Jumlah Penduduk dan Kepadatan Penduduk memiliki angka 0.614 angka ini menunjukkan adanya korelasi yang cukup kuat dan positif atau searah. Ini berarti keterhubungan masih cukup untuk dilakukan pengujian analisis kluster karena saling terarah.

2.3 Interpretasi Analisis Klaster

penelitian ini dilakukan untuk clustering jenis penyakit menular pada Puskesmas Kota Tangerang dengan data dari BPS Kota Tangerang, dokumen publikasi dengan nama Katalog Kecamatan dalam Angka 2018 pada bulan April-Mei 2019. Data 13 Kecamatan 30 Puskesmas yang akan di klusterkan menjadi empat Klaster dengan Klaster 1 : TB Paru Klaster 2 : Pneumonia Klaster 3 : DBD Klaster 4 : Diare.

Untuk itu agar dapat terbentuk klaster yang diinginkan akan diujikan memakai aplikasi SPSS 22 dengan interpretasi analisis klaster menggunakan metode K-means Cluster yang akan dibahas di bawah ini :

Tabel 2.3 *Descriptive Statistics*

	N	Minimum	Maximum	Mean	Std. Deviation
Jumlah Penduduk	30	27103	198950	86157.53	44464.948
Kepadatan Penduduk	30	2657	171267	60970.87	40789.569
TB.Paru	30	9	97	38.13	20.384
Pneumonia	30	55	708	198.10	127.336
DBD	30	0	41	14.50	10.951
Diare	30	290	1762	962.70	427.916
Valid N (listwise)	30				

Keluaran tabel ini untuk membantu pengelompokan klaster dengan Zscore yang dihasilkan pada uji ini. Sehingga dapat menganalisa kluster yang terbentuk nantinya melalui Zscore.

Tabel 2.4 *Initial Cluster Centers*

	Cluster			
	1	2	3	4
Zscore: Jumlah Penduduk	-1.25381	2.53666	.85255	-.37280
Zscore: Kepadatan Penduduk	.72894	2.70403	.02820	-1.41058
Zscore(TB.Paru)	-1.42924	-.15372	.04252	.48404
Zscore(Pneumonia)	-1.12380	-.27565	2.06462	-.64475
Zscore(DBD)	-.50226	-1.05018	1.96338	.13698
Zscore(Diare)	-1.57204	1.11774	-.98080	1.86789

Pada tabel initial cluster centers ini merupakan proses pertama untuk pengelompokan data aplikasi SPSS 22. Karena pengelompokan data ini merupakan proses untuk pembentukan empat kluster secara kasarnya untuk melihat pembentukan klaster lihat pada tabel berikut:

Tabel 2.5 Iteration History

Iteration	Change in Cluster Centers			
	1	2	3	4
1	1.826	.893	1.460	1.542
2	.183	.000	.000	.211
3	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The current iteration is 3. The minimum distance between initial centers is 4.635.

Tabel di atas adalah tabel *iteration history* yaitu proses pengelompokan kluster melalui iterasi untuk membaca tabel ini, karena menggunakan *K-means Cluster* tabel ini menjelaskan bahwa melalui proses iterasi dengan gambaran tabel iterasi di atas iterasi terjadi 3 kali, karena pada iterasi 1 = *centeroid* tidak signifikan iterasi 2 = nilai *centeroid* hampir signifikan iterasi 3 = *centeroid* berada dalam signifikan semua maka kluster sudah terbentuk dan iterasi stop. dengan kluster yang telah ditentukan yaitu empat kluster memiliki proses iterasi sebanyak 3 kali iterasi dengan jarak minimum 4.635.

3. HASIL DAN DISKUSI

3.1 Validasi Analisis Kluster

Validasi analisis kluster ini dapat dilihat pada tabel 3.1 *cluster membership* dari interpretasi yang tafsirkan bahwa kluster terbentuk melalui tiga kali iterasi, disini akan dibahas hasil validasi dari iterasi tersebut sehingga terbentuk kluster pada tabel berikut ini :

Tabel 3.1 Cluster Membership

Cluster Membership			Cluster Membership		
Case Number	Cluster	Distance	Case Number	Cluster	Distance
1	4	.956	16	4	2.082
2	4	1.287	17	1	1.904
3	4	1.578	18	2	.893
4	1	1.332	19	4	2.169
5	3	1.687	20	4	1.615
6	3	2.881	21	4	1.517
7	1	1.318	22	4	1.582
8	1	2.536	23	1	1.930
9	1	2.008	24	4	1.509
10	3	2.283	25	4	1.849
11	3	1.460	26	1	1.288
12	1	.932	27	2	.893
13	3	3.059	28	1	1.491
14	1	2.498	29	4	1.272
15	1	1.150	30	1	1.567

Keluaran dari tabel 3.1 Cluster Membership dapat dikatakan hasil akhir kluster dengan aplikasi SPSS 22, yang menggunakan iterasi sebanyak tiga kali iterasi.

3.2 Profiling Analisis Klaster

Dari hasil validasi yang dapat dilihat pada tabel 3.1 selanjutnya adalah Profiling sehingga dapat dilihat untuk klaster pertama terbagi menjadi 12 Puskesmas, klaster kedua terbagi dalam 2 Puskesmas, klaster ketiga ada pada 5 Puskesmas, dan untuk klaster keempat terbagi pada 11 Puskesmas, dimana terjadi pengelompokan di Puskesmas Perkecamatan masing-masing klaster yang dijelaskan pada tabel dibawah ini:

Tabel 3.2 Hasil Akhir Klaster Pada Puskesmas

Data ke-i	Klaster 1	Klaster 2	Klaster 3	Klaster 4
1				Batu Ceper
2				Poris Gaga
3				Jurumudi Baru
4	Benda			
5			Cibodasari	
6			Baja	
7	Tajur			
8	Ciledug			
9	Gondrong			
10			Ketapang	
11			Jatiuwung	
12	Pondok Bahar			
13			Karang Tengah	
14	Pabuaran Tumpeng			
15	Bugel			
16				Karawaci Baru
17	Pasar Baru			
18		Larangan Utara		
19				Cipadu
20				Neglasari
21				Kedaung Wetan
Data ke-i	Klaster 1	Klaster 2	Klaster 3	Klaster 4
22				Periuk Jaya
23	Sangiang Jaya			
24				Gembor
25				Kunciran
26	Kunciran Baru			
27		Panungangan		
28	Sukasari			
29				Tanah Tinggi
30	Cikokol			

Tabel 3.2 menunjukkan Hasil Akhir Klaster Pada Puskesmas dan dapat diartikan dengan profiling sebagai berikut :

1. Klaster 1

TB Paru adalah penamaan untuk klaster 1 maka dari itu klaster 1 hasil akhir pengklasteran mendapatkan 12 Puskesmas dengan nama daerah puskesmas yang terbagi klaster 1 yaitu : Puskesmas Benda memiliki jarak 1.332, Puskesmas Tajur memiliki jarak 1.318, Puskesmas ciledug memiliki jarak 2.536, Puskesmas Gondrong memiliki jarak 2.008, Puskesmas Pondok Bahar memiliki jarak 0.932, Puskesmas Pabuaran Tumpeng memiliki jarak 2.498, Puskesmas Bugel memiliki jarak 1.150, Puskesmas Pasar Baru memiliki jarak 1.904, Puskesmas Sangiang Jaya memiliki jarak 1.930, Puskesmas Kunciran Baru memiliki jarak 1.288, Puskesmas Sukasari memiliki jarak 1.491, dan Puskesmas Cikokol memiliki jarak 1.567. dengan jarak pada masing-masing klaster yang paling dominan adalah Puskesmas Ciledug dengan jarak antar titik memiliki 2.536 dengan jumlah penderita TB Paru sebanyak 75 orang, jumlah penduduk 87683, kepadatan penduduk 91563. Untuk itu dengan kepadatan penduduknya melebihi jumlah pada suatu daerah artinya penduduk tersebut memiliki kepadatan pada daerahnya sehingga ruang untuk penyebaran penyakit TB Paru ini dapat mudah menular. dari masing-masing puskesmas yang terdapat pada klaster 1 puskesmas ciledug lah yang memiliki paling banyak menderita penyakit TB Paru ini.

2. Klaster 2

Pneumonia adalah penamaan untuk klaster 2 maka dari itu klaster 2 hasil akhir pengklasteran mendapatkan 2 Puskesmas saja, dengan nama daerah puskesmas yang terbagi klaster 2 yaitu : Puskesmas Larangan Utara memiliki jarak 0.893, dan Puskesmas Panunggangan memiliki jarak 0.893, dengan jarak yang sama sehingga karakteristik kedua puskesmas sama tetapi memiliki kasus penderita pneumonia yang berbeda. Untuk Puskesmas Panunggangan ada 180 kasus dan Puskesmas Larangan Utara 165 kasus, sehingga pada klaster 2 Puskesmas yang paling dominan adalah Puskesmas Panunggangan dengan 180 kasus.

3. Klaster 3

DBD adalah penamaan untuk klaster 3 maka dari itu klaster 3 hasil akhir pengklasteran mendapatkan 5 Puskesmas dengan nama daerah puskesmas yang terbagi klaster 3 yaitu : Puskesmas Cibodasari memiliki jarak 1.687, Puskesmas Baja jarak 2.881, Puskesmas Ketapang memiliki jarak 2.283, Puskesmas Jatiuwung memiliki jarak 1.460, dan Puskesmas Karang Tengah memiliki jarak 3.059. dengan jarak pada masing-masing klaster yang paling dominan adalah Puskesmas Ketapang dengan jarak antar titik memiliki 2.283 dengan jumlah penderita DBD sebanyak 41 orang, jumlah penduduk 155801, kepadatan penduduk 50819. dengan jumlah yang terjangkau penyakit DBD lebih banyak dari puskesmas lainnya yang berada di klaster 3 yaitu berjumlah 41 orang.

4. Klaster 4

Diare adalah penamaan untuk klaster 4 maka dari itu klaster 4 hasil akhir pengklasteran mendapatkan 11 Puskesmas dengan nama daerah puskesmas yang terbagi klaster 4 yaitu : Puskesmas Batu Ceper memiliki jarak 0.956, Puskesmas Poris Gaga memiliki jarak 1.287, Puskesmas Jurumudi Baru memiliki jarak 1.578, Puskesmas Karawaci Baru memiliki jarak 2.082, Puskesmas Cipadu memiliki jarak 2.169, Puskesmas Neglasari memiliki jarak 1.615, Puskesmas Kedaung Wetan memiliki jarak 1.517, Puskesmas Periuk Jaya memiliki jarak 1.582, Puskesmas Gembor memiliki jarak 1.509, Puskesmas Kunciran memiliki jarak 1.849, dan Puskesmas Tanah Tinggi memiliki jarak 1.272. dengan jarak pada masing-masing klaster yang paling dominan adalah Puskesmas gembor dengan jarak antar titik memiliki 1.509

dengan jumlah penderita Diare sebanyak 1762 orang, jumlah penduduk 69581, kepadatan penduduk 3434.

Untuk Lebih Detail lagi keluaran tabel selanjutnya akan membahas centeroid pada Hasil klaster yang telah dibentuk pada aplikasi SPSS 22 berikut adalah tabel 4.9 Final Cluster Centers:

Tabel 3.3 Final Cluster Centers

	Cluster			
	1	2	3	4
Zscore: Jumlah Penduduk	-.28056	2.40152	.89789	-.53871
Zscore: Kepadatan Penduduk	.01642	1.93772	.43090	-.56609
Zscore(TB.Paru)	-.42763	.01799	.74896	.12280
Zscore(Pneumonia)	-.55051	-.20890	1.52746	-.05576
Zscore(DBD)	.00000	-.82188	1.01365	-.31132
Zscore(Diare)	-.80612	.79174	-.45266	.94120

Pembacaan tabel diatas merupakan proses akhir dari pengelompokan yang dilakukan, yaitu sebanyak empat klaster yang terbentuk untuk masing-masing variabel. Angka-angka yang dihasilkan terbentuk Titik Center pada Zscore yang telah diketahui sebelumnya dengan hasil akhir centeroid pada tabel diatas, untuk dapat membaca hasil akhir profiling antar variabel yang akan dibahas disini adalah hanya variabel Jumlah Penduduk dan Kepadatan Penduduk. Pengaruh kedua variabel tersebut dengan keempat klaster pada pengelompokan ini, dapat menggunakan rumus sebagai berikut:

$$X = \mu + z \cdot \sigma \quad (3.1)$$

Keterangan :

X = rata-rata sampel/data/variabel dalam klaster tertentu

μ = rata-rata populasi

σ = standar deviasi

z = nilai standarisasi

Berikut rata-rata populasi klaster 1,2,3, dan 4 dari variabel Jumlah Penduduk:

1. Klaster 1 :

$$\text{Rata-rata Variabel 1} + (\text{nilai } z \text{ klaster 1} \times \text{standar deviasi rata-rata variabel 1}) = 86157,35 + (-0,28056 \times 44464,948) = 73682,444$$

2. Klaster 2

$$\text{Rata-rata Variabel 1} + (\text{nilai } z \text{ klaster 2} \times \text{standar deviasi rata-rata variabel 1}) = 86157,35 + (2,40152 \times 44464,948) = 192940.991$$

3. Klaster 3

$$\text{Rata-rata Variabel 1} + (\text{nilai } z \text{ klaster 3} \times \text{standar deviasi rata-rata variabel 1}) = 86157,35 + (0,89789 \times 44464,948) = 126082.162$$

4. Klaster 4

$$\text{Rata-rata Variabel 1} + (\text{nilai } z \text{ klaster 4} \times \text{standar deviasi rata-rata variabel 1}) = 86157,35 + (-0,53187 \times 44464,948) = 62203,818$$

Dari perhitungan yang di dapat bahwa Variabel Jumlah Penduduk terdapat empat klaster. Klaster 1 sebesar 73682,444 berarti di bawah rata-rata populasi; untuk klaster 2 192940,991 di atas rata-rata populasi; untuk klaster 3 sebesar 126082,162 di atas rata-rata populasi; dan klaster ke 4 sebesar 62203,818 dibawah rata-rata populasi. Rata-rata populasinya adalah 86175,53 (lihat di tabel Descriptive Statistics). Maka Penduduk yang memiliki di bawah rata-rata populasi atau di bilang jumlah di bawah rata-rata mendapat pengelompokan jenis penyakit menular TB Paru dan diare, sedangkan penduduk yang di atas rata-rata populasi atau dapat dikatakan jumlah penduduk di atas rata-rata mendapat pengelompokan jenis penyakit menular Pneumonia dan DBD.

Berikut rata-rata populasi klaster 1,2,3, dan 4 dari variabel Kepadatan Penduduk:

1. Klaster 1 :

$$\text{Rata-rata Variabel 2} + (\text{nilai } z \text{ klaster 1} \times \text{standar deviasi rata-rata variabel 2}) = 60970,87 + (0,1642 \times 40789,569) = 67668,517$$

2. Klaster 2

$$\text{Rata-rata Variabel 2} + (\text{nilai } z \text{ klaster 2} \times \text{standar deviasi rata-rata variabel 2}) = 60970,87 + (1,93772 \times 40789,569) = 140009,634$$

3. Klaster 3

$$\text{Rata-rata Variabel 2} + (\text{nilai } z \text{ klaster 3} \times \text{standar deviasi rata-rata variabel 2}) = 60970,87 + (0,43090 \times 40789,569) = 78547,095$$

4. Klaster 4

$$\text{Rata-rata Variabel 2} + (\text{nilai } z \text{ klaster 4} \times \text{standar deviasi rata-rata variabel 2}) = 60970,87 + (-0,56609 \times 40789,569) = 37880,303$$

Dari perhitungan yang di dapat bahwa Variabel Kepadatan Penduduk terdapat empat klaster. Klaster 1 sebesar 67668,517 berarti di atas rata-rata populasi; untuk klaster 2 140009,634 di atas rata-rata populasi; untuk klaster 3 78547,095 sebesar di atas rata-rata populasi; dan klaster ke 4 sebesar 37880,303 dibawah rata-rata populasi. Rata-rata populasinya adalah 60970,87 (lihat di tabel Descriptive Statistics). Maka Penduduk yang memiliki di bawah rata-rata populasi atau di bilang Kepadatan Penduduk di bawah rata-rata mendapat pengelompokan jenis penyakit menular diare, sedangkan penduduk yang di atas rata-rata populasi atau dapat dikatakan Kepadatan penduduk di atas rata-rata mendapat pengelompokan jenis penyakit menular TB Paru, Pneumonia, dan DBD.

3.3 Asumsi-asumsi Analisis Klaster

Asumsi-asumsi pada analisis multivariate biasanya menggunakan uji normalitas, uji linearitas, dan uji heteroskedastisitas. Untuk asumsi pada Analisis Klaster biasanya memakai uji ANOVA untuk melakukan pengujian secara signifikan antar klaster. Berikut tabel 4.10 ANOVA:

Tabel 3.4. ANOVA

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Zscore: Jumlah Penduduk	6.567	3	.358	26	18.366	.000
Zscore: Kepadatan Penduduk	3.989	3	.655	26	6.088	.003
Zscore(TB.Paru)	1.722	3	.917	26	1.878	.158
Zscore(Pneumonia)	5.141	3	.522	26	9.846	.000
Zscore(DBD)	2.518	3	.825	26	3.053	.046
Zscore(Diare)	6.607	3	.353	26	18.713	.000

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

Tabel ANOVA digunakan untuk melihat tingkat signifikansi antar klaster. Kolom klaster merupakan besaran *between cluster means* dan kolom Error merupakan *within cluster means*. Sementara itu, nilai kolom F diperoleh melalui rumus berikut :

$$F = \frac{\text{between means}}{\text{within means}}$$

Ketentuan penggunaan angka F dalam analisis klaster ialah bahwa semakin besar angka F hitung (jika dilakukan uji Hipotesis, maka F hitung akan lebih besar dari F tabel) dan tingkat signifikansi ($\text{sig} < 0,05$); maka semakin besar perbedaan antara ketiga klaster yang terbentuk.

Untuk itu peneliti akan membuat hipotesis pengujian sebagai berikut :

Merumuskan Hipotesis :

H_0 : Keempat klaster tidak mempunyai perbedaan yang signifikan.

H_1 : Keempat klaster mempunyai perbedaan yang signifikan.

Jika angka signifikan $>0,05$; H_0 diterima H_1 ditolak

Jika angka signifikan $<0,05$; H_0 ditolak H_1 diterima

$$F = \frac{18,366}{0,358} = 51,30$$

Jika dilihat dari perhitungan variabel 1 hasil F hitung $> F$ tabel maka dengan angka 51,30 $> 18,366$ dan signifikansi 0,000; maka H_0 ditolak H_1 diterima. Artinya, keempat klaster mempunyai perbedaan yang signifikan. Dan dari semua variabel hampir mempunyai signifikansi $<0,05$ maka dari keseluruhan dapat dikatakan mempunyai perbedaan yang signifikan.

Dengan demikian hasil klaster yang di dapat dalam penelitian ini bahwa untuk pengelompokan jenis penyakit menular menunjukkan adanya perbedaan pada keempat klaster yang terbentuk. Hal ini dengan ditunjukkannya nilai $F = 52,30$ dan $\text{sig} = 0,001$. Selanjutnya untuk mengetahui jumlah anggota masing-masing klaster yang terbentuk dapat dilihat pada tabel berikut ini:

Tabel 4.11 Number of Cases in Each Cluster

Cluster	1	12.000
	2	2.000
	3	5.000
	4	11.000
Valid		30.000
Missing		.000

Nampak jelas bahwa kluster 1 beranggotakan 12 Puskesmas, kluster 2 beranggotakan 2 Puskesmas, kluster 3 beranggotakan 5 Puskesmas dan untuk kluster 4 beranggotakan 11 Puskesmas.

4. KESIMPULAN

Penelitian ini menghasilkan empat kluster dengan profiling kluster 1 TB Paru kluster 2 Pneumonia kluster 3 DBD kluster 4 Diare. Proses terbentuknya kluster-kluster tersebut adalah dengan iterasi-iterasi yang dilakukan, pada penentuan kluster ini terjadi 3 kali iterasi yang dilakukan sehingga mendapat hasil nilai minimum yang sampai terbentuk keempat kluster tersebut, dengan nilai minimum akhir 4.635. TB Paru adalah penamaan untuk kluster 1 maka dari itu kluster 1 hasil akhir pengklasteran mendapatkan 12 Puskesmas, Pneumonia adalah penamaan untuk kluster 2 maka dari itu kluster 2 hasil akhir pengklasteran mendapatkan 2 Puskesmas, DBD adalah penamaan untuk kluster 3 maka dari itu kluster 3 hasil akhir pengklasteran mendapatkan 5 Puskesmas, Diare adalah penamaan untuk kluster 4 maka dari itu kluster 4 hasil akhir pengklasteran mendapatkan 11 Puskesmas. 3. Kluster yang paling dominan dari kluster yang terbentuk adalah Kluster 1.

5. REFERENSI

- Hasan, M. I. (2002). Pokok-pokok Metodologi Penelitian dan Aplikasinya. Jakarta: Ghalia Indonesia.
- Martono, N. (2011). Metode Penelitian Kuantitatif: Analisis Isi dan Analisis Data Sekunder. Jakarta: PT.Raja Grafindo Persada.
- Prof. Dr. Sugiono. (2013). Statistika Dalam Penelitian. Bandung: ALFABETA.
- Prof. Gundono, Ph.D. CMA. (2017). Analisis Data Multivariat. Yogyakarta: BPF.
- Sarwono, J. (2015). Rumus-rumus Populer dalam SPSS 22 untuk Riset Skripsi. Yogyakarta: ANDI.
- Sugiono. (2009). Penelitian Kuantitatif Kualitatif dan R&D. Bandung: ALFABETA.
- Suharsimi. (2002). Metodologi Penelitian. Yogyakarta: ALFABETA.

PUSTAKA DARI INTERNET

- B.J, P. (2008, Desember Senin). Analisis Cluster. Retrieved Mei Minggu, 2019, from prayudho.wordpress.com: <https://prayudho.wordpress.com/2008/12/30/analisis-cluster/>

- Badan Pusat Statistik kota Tangerang. (2017). Statistik daerah kota Tangerang. Kota Tangerang: CV.Dharma Putra.
- depkes.go.id. (2004). Sejarah Puskesmas. Depkes.
- Dr. Kevin Adrian. (2018, September Jumat). Penyakit menular yang umum di indonesia. Retrieved April Senin, 2019, from alodokter: <http://www.alodokter.com>
- Dr. Tjin Willy. (2018, Desember Jumat). Pneumonia. Retrieved April Senin, 2019, from alodokter: <http://www.alodokter.com>
- Hidayat, A. (2014, Maret). Statistikian. Retrieved Februari Senin, 2019, from www.statistikian.com: <https://www.statistikian.com/2014/03/analisis-cluster.html/amp>
- Kementrian kesehatan. (2010, September Rabu). PusDatin. Retrieved May Senin, 2019, from www.pusdatin.kemkes.go.id: <http://www.pusdatin.kemkes.go.id>
- Kementrian Kesehatan RI. (2015). Hari Tuberkulosis Sedunia. Jakarta: InfoDATIN.